

THE NEURAL NETWORK AS A RENORMALIZER OF INFORMATION

BY

WILLARD L. MIRANKER

Department of Computer Science, Yale University, New Haven, CT

Abstract. We characterize the behavior of information in neural processing as the neuronal circuitry itself agglomerates into assemblies of increasing size and complexity. The basic synaptic stage of this processing is interpreted as the observer feature of a measurement process, a quality that extends up the assembly hierarchy. Renormalization techniques are employed, and they supply features of emergence to the information. Renormalization also supplies each observer feature with a measurable physical quantity called a token, the latter supplying quantitative aspects to the entire development. This development is used to frame an analytic theory of phenomenal consciousness, featuring emergent aspects. *The tokens furnish the means for the various predictions and explanations of that theory to be subjected to measurement and experimental verification.*

1. Introduction. The information processed dynamically by neurons as the latter agglomerate into a hierarchy of assemblies is characterized, employing a representation of that processing as a collection of measurements. The basic synaptic stage is interpreted as the observer feature of such a measurement, a quality that extends upward to each level of the hierarchy. To develop this characterization we provide extensions (adaptations) of two developments from classical physics: the Ising model and the renormalization theory.

The Ising model finds its development in the equilibrium analysis of magnetic spins. It has inspired an extensive body of work dealing with neural networks ([1], [2], [22], [23], [33]) in which spins of the Ising model become the bits that define the information processed by neural circuitry. We extend these Ising model methods by means of a statistical mechanics-like study of the *Hebbian dynamics* for synaptic weight change.

The Renormalization Group technique (as applied to magnetic spin phenomena) employs a hierarchy of coarsening into blocks of spins (an agglomerating or ramification methodology) of the regular lattice structure of a ground model ([25], [48], [14]...). Although the neocortex consists of layers and columns, its connectivity is far from the

Received January 4, 2007.

2000 *Mathematics Subject Classification.* Primary 91E30, 46N60.

Key words and phrases. Consciousness, dualist construct, emergence, Hebbian dynamics, measurement, neural networks, phase change, renormalization.

©2008 Brown University
Reverts to public domain 28 years from publication

regularity needed to attribute a lattice-like structure to it. So *we extend the renormalization methodology to a hierarchy of neuronal assemblies, namely to structures of arbitrary geometry and with an arbitrary specification of proximity within those structures. This demonstrates a novel validity of renormalization for appropriate irregular physical systems generally, regularly structured or not.* The renormalization approach supports development of what we call a token¹ of observation at each level of the neuronal assembly hierarchy. A token is specified by an analytic formula that expresses it in terms of measurable physical quantities.

We employ these two tools of classical physics (as extended) to model emergence features of information in neural networks. We derive the renormalization transformation that in classical physics leads to the phase diagrams that characterize the different phases of matter. Our generalizations convey the corresponding analytical methods to types of structures found in biology. (See [3], [15].)

1.1. Application to the study of consciousness. We apply the extensions developed to frame a quantitative and therefore testable theory of consciousness. Our approach is motivated by observing that dualist constructs are critical constituents of scientific methodology ([29], [30]). We formulate such a construct at the level of neurons and their synapses. Hebbian synaptic dynamics characterized as an observer process of the act of making a measurement (of the correlation of a neuron's afferents and its efferent) is taken as a primitive awareness quality, i.e., as a dualist construct embodying a synaptic level of (proto-) consciousness. At each level of the hierarchy the renormalization specifies an awareness quality and a corresponding token, the value of the latter supplying a quantitative measure of the awareness. We show that *there is a magnification of the value of the associated tokens as the neuronal assembly hierarchy is ascended, and this feature forms the basis for experimental verification of the aspects of the descriptors of consciousness developed.*

1.2. Outline. In Section 2 a neural network model is introduced, and its activity at the synapse-neuron level is identified as the observer aspect of a measurement process. In Section 3 we briefly review the statistical mechanics methodology applied to the neural network in the presence of noise as a basis for what follows. In Section 4 we extend that application to characterize Hebbian synaptic dynamics as an observer aspect of the measurement process, and we develop a token of observation of the measurement that turns out to be the neural network (the information) analog of the spin correlation function in the physics of magnetism. In Section 5 we use a construct called a *lattice brain*. We review (briefly for expository purposes) the use of the renormalization methodology for block spins in a lattice, and we reinterpret it in the context of the lattice brain as a way to ramify the synaptic level of measurement (in particular, the observation quality) to observation at higher levels of neural network organization. In Section 6 we dispense with the need for lattice regularity by extending validity of the renormalization technique to a fully general hierarchy of neuronal assemblies (that is, one with arbitrary notions of

¹A token of a qualitative property is a quantitative measure of that property. For example, using our terminology, we would say that the pressure p is a token of the specific force, the pressure being a measurable quantity, while the quality of specific force is a dualist construct of Newtonian mechanics.

structure and proximity). There results both the equation that specifies the renormalization transformation in the fully general case as well as an analytic expression for what we call the observation-quality/observation-token at the assembly level. In Section 7 we apply the developments of Sections 2–6 to frame a quantitative theory of consciousness. We reinterpret the synaptic level observer aspect as a primitive awareness quality and append it to the development as a dualist construct. A number of interpretations of features of consciousness are then formulated in terms of the developments of Sections 2–6. *The use of the tokens for experimental verification of both the theory and these associated interpretations is described.* Some of the analysis used is known, so many details are suppressed with references supplied.

2. Neural dynamics as a measurement process.

2.1. Network of McCulloch-Pitts neurons, Hebbian synaptic dynamics. For convenience and clarity we take the information flow in a network of N neurons to be described by the following McCulloch-Pitts dynamics ([31], [16]).

$$S_i(t+1) = \text{sgn}\left(\sum_{j=1}^N w_{ij} S_j(t) + h_i^{ex}\right), i = 1, \dots, N, \quad \text{sgn}(x) = \begin{cases} 1, & x \geq 0, \\ -1, & x < 0, \end{cases} \quad (2.1)$$

where S_i is the output of neuron i . Hereafter, we shall omit the specification of index ranges when the meaning is clear. The synaptic connections from each of the N neurons to neuron i are characterized by the vector of synaptic weights $w_i = (w_{i1}, \dots, w_{iN})$. h_i^{ex} is the exogenous input to neuron i .

Information (e.g., a memory trace) is taken as encoded (associatively) in the neural net by adjustments to its synaptic weights. As the neurons process information according to (2.1), the synaptic weights change according to Hebbian synaptic dynamics,

$$\frac{dw_{ij}(t)}{dt} = \mathcal{H}(S_i(t), S_j(t-1)), \quad (2.2)$$

where \mathcal{H} is the so-called Hebb function. Hebb's proposal may be expressed by saying that \mathcal{H} is to have the sign of the correlation of its two arguments ([17]). One specification of \mathcal{H} that achieves this is made in (4.1). Note that our description of Hebb's idea is not limited to the case that neurons firing together increase their interaction strength. There have been many numerical studies of Hebb's assembly theory (exemplified by (2.1), (2.2)). (See [11], [12].)

2.2. The measurement/observation process, token of observation. Referring to Figure 2.1, we characterize the interrelated dynamical systems (2.1)–(2.2) as a measurement/observation process. The input is the vector $S(t) = (S_1(t), \dots, S_N(t))$. The measuring device is the neuron. The value of the measurement is the vector of neuronal outputs $S(t+1)$ specified in (2.1). The *observer of the measurement is the biochemical process* that generates the Hebbian synaptic dynamics (2.2).

Discretizing and indexing time instants with n , the Hebbian observation process is described in terms of the change Δw_{ij} in synaptic weight:

$$\Delta w_{ij}(n) = \mathcal{H}(S_i(n), S_j(n-1)). \quad (2.3)$$

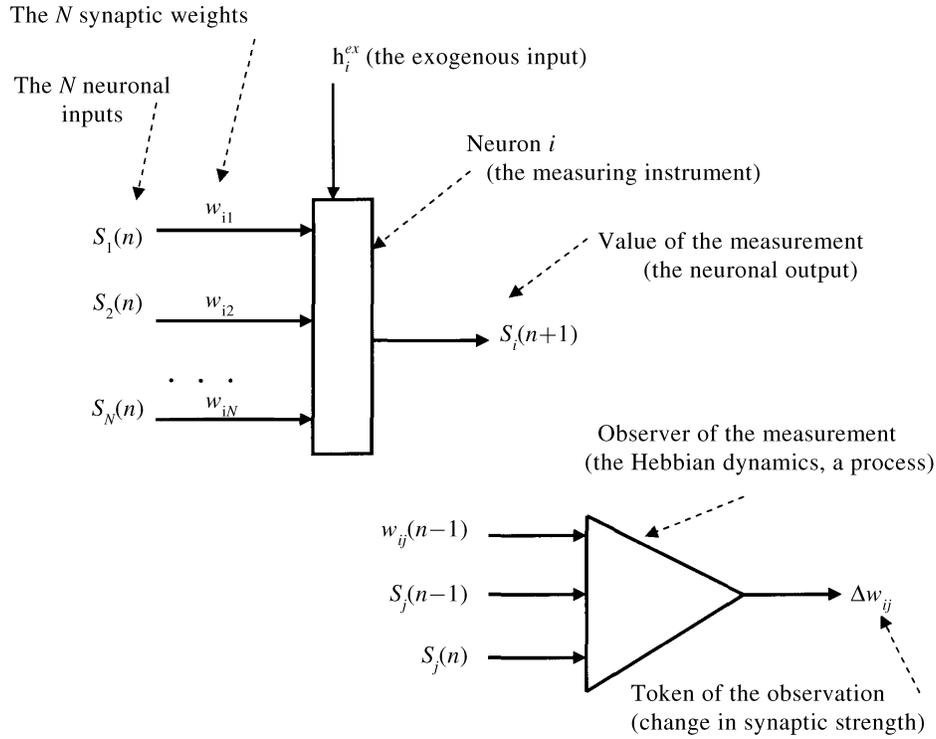


FIG. 2.1. The neural input/output dynamics (upper left) and the Hebbian synaptic dynamics (lower right) showing the token of observation

This $\Delta w_{ij}(n)$ is a quantitative measure of the synaptic level observer process that we shall refer to as the *token of observation*. That is, an expression for the magnitude of the observation associated with the weight change dynamics of the j -th input synapse of neuron i .

3. Neural dynamics (measurement) in the presence of noise. A brief review is given of developments illustrating change of phase for information and for forming a basis for study in Section 4 of the measurement process associated with the dynamics in Section 2, all this in the presence of noise (see [22] for details). In the presence of noise, the neuronal output (2.1) is modeled by the stochastic output rule

$$Pr(S_i = \pm 1) = \frac{1}{1 + \exp(\mp 2\beta h_i)}. \quad (3.1)$$

Here h_i denotes the total input to neuron i , and $\beta = 1/T$, where T is a (pseudo-) temperature, a parameter controlling the noise level.

The exogenous input h_i^{ex} to neuron i may be viewed as coming from another neural network, perhaps that of a sense organ. Now suppose that the neural net has been used to store a set of p so-called fundamental memories, $\xi^\mu = (\xi_1^\mu, \dots, \xi_N^\mu)$, $\mu = 1, \dots, p$. This means that such a neural net, when cued by an input h_i^{ex} close enough to ξ_i^μ , will produce

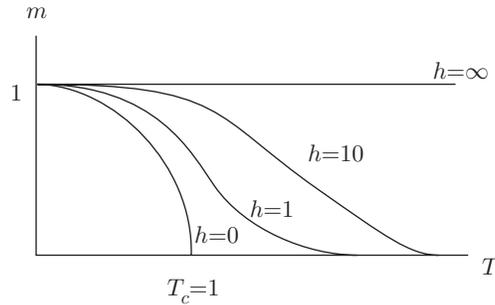


FIG. 3.1. A plot of $m(h, T)$ demonstrating phase change in the retrieval process

$S_i = \xi_i^\mu$ as a steady state output. One way to assure this is to specify the synaptic weights as $w_{ij} = \frac{1}{N} \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu$ (see chap. 2 of [16]).

3.1. *Mean field approximation.* The mean field approximation for the total input to the i -th neuron is

$$\langle h_i \rangle = \sum_j w_{ij} \langle S_j \rangle + h_i^{ex}, \tag{3.2}$$

where $\langle \rangle$ denotes the expected value. Suppose the exogenous input is proportional to a stored pattern $h_i^{ex} = h \xi_i^v$, where h is a constant (e.g., suppose that the exogenous signal comes from a sense organ and closely resembles an earlier such recorded signal, say, the fundamental memory ξ^v). This motivates the hypothesis that $\langle S_i \rangle$ is proportional to the same stored pattern. In particular, for some constant m , we set

$$\langle S_i \rangle = m \xi_i^v = \pm m. \tag{3.3}$$

Then it is known that

$$m = \tanh h\beta(m + h). \tag{3.4}$$

Then the expected number of correct outputs of the neural net (the expected number of bits in the fundamental memory ξ^v that are recalled correctly) in response to an arbitrary cue is

$$\langle N_{\text{correct}} \rangle = \frac{1}{2} N(1 + m). \tag{3.5}$$

(For a derivation of (3.2) see p.29 of [22]. For (3.4) and (3.5) see p. 34.)

3.2. *A phase change of information.* In Figure 3.1 we plot $m = m(h)$ versus T given by (3.4) for different values of h . The following observations flow from that figure. For $h = 0$, we see a change of phase effect. Namely for T greater than a critical temperature, $T_c = 1$, we have $m = 0$. Then from (3.5), the expected number of correct bits in the retrieved memory is $N/2$, a random result. As T falls below T_c and approaches zero (the no-noise limit), we see a rapid rise of m to unity, i.e., a rapid rise to full correctness of the expected number of retrieved bits. For any $T \geq 0$, the expected number of correctly retrieved bits increases with the strength h of the exogenous input, and it approaches N (full correctness) as the exogenous signal strength $h \rightarrow \infty$.

4. The token of observation in the presence of noise.

4.1. *Expected value of the token of observation.* A simple form for $\mathcal{H}(x, y)$ that characterizes the correlation properties of synaptic weight development enunciated by Hebb (see (2.2)f) is the product form (for details see p.57 of [16]), namely

$$\mathcal{H}(x, y) = \kappa(x - \langle x \rangle)(y - \langle y \rangle), \quad (4.1)$$

where κ may be taken as unity without loss of generality. Combining (4.1) with (2.3) and taking expected values, a direct averaging calculation shows that

$$\langle \Delta w_{ij} \rangle = \langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle. \quad (4.2)$$

So the expected value of the token of observation $\langle \Delta w_{ij} \rangle$ (see (2.3)) is proportional to the neural output correlation function.² According to our terminology, we shall call the left member of (4.2) the expected value of observation (dropping the term token), since confusion should not result. (We shall do this for other types of observation tokens as well.) This motivates the following.

DEFINITION. *During a step of neural processing, the ij -th synapse makes the observation σ_{ij} , where*

$$\sigma_{ij} = \langle \Delta w_{ij} \rangle. \quad (4.3)$$

We stress that while both symbols in (4.3) are tokens, observation itself characterizes the behavior of the ij -th synapse (indeed the entire Hebbian synaptic dynamics process) as an observer of what we call the neuronal measurement process. (See Figure 2.1.)

4.2. *The cellular level observation token, mirroring.* The average over a neuron's synaptic observation tokens

$$\mathcal{I}_i = \frac{1}{N} \sum_j \sigma_{ij} \quad (4.4)$$

will be a derived cellular level token of interest. Let

$$\bar{S} = \frac{1}{N} \sum_j S_j \quad (4.5)$$

denote the average output of the neural network. (In the case of modeling magnetic spins, $\text{sgn}(\bar{S})$ is referred to as the block spin.) Using (4.2)–(4.5), we find

$$\begin{aligned} \mathcal{I}_i &= \frac{1}{N} \sum_j [\langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle] \\ &= \langle \bar{S} S_i \rangle - \bar{S} \langle S_i \rangle. \end{aligned} \quad (4.6)$$

Setting $\bar{\bar{S}} = \langle \bar{S} S_i \rangle / \langle S_i \rangle$, (4.6) simplifies, becoming

$$\mathcal{I}_i = (\langle \bar{\bar{S}} \rangle - \bar{S}) \langle S_i \rangle. \quad (4.7)$$

Combining this with (3.3), we find

$$\mathcal{I}_i = \pm (\langle \bar{\bar{S}} \rangle - \bar{S}) m. \quad (4.8)$$

Thus \mathcal{I}_i is proportional to m , the expected value of neuronal output $\langle S_i \rangle$. So the subject of the cellular level observation represented by the token \mathcal{I}_i is the neuronal output itself.

²In the modeling of magnetic spins S_i , the right member of (4.2) is called the spin correlation function.

We call this a *mirroring of the expected value of the neuronal output* (a measurable quantity) by the token \mathcal{I}_i .

4.3. *Neural dynamics as a communication process, mutual information.* Since many different patterns of input synaptic activity can correspond to each one of the two possible neuronal output values, $S_i = \pm 1$, all we can expect to tell about the input pattern from the value of the output is the average (counting signatures) of the inputs. The mirroring expressed by (4.8) tells us this. Then \mathcal{I}_i is the *mutual information* of the neuronal input/output dynamics (subject to noise). (Mutual information is the uncertainty resolved about an input by knowledge of the output. See Section 11.2 of [16] for a discussion of the concepts of information theory in the context of neural networks.)

4.4. *Neuronal level observation and token.* According to the definition in Section 4.1, \mathcal{I}_i is an appropriate average (having either sign) of the tokens of observation σ_{ij} of the input synapses of neuron i (see (4.3), (4.4)). Then we say that \mathcal{I}_i is a token of observation of the entire neuron, which we shall for convenience refer to as the observation quality of the neuron. (We stress once more the need to keep in mind the difference between a qualitative process and its token.) The \pm sign in (4.8) suggests that neuronal observation (that is, the mutual information of the neural measurement process) may be attractive or repulsive. Indeed \mathcal{I}_i is expressed in terms of the measurable neuronal output activity (see (4.8)f), that activity encoding the neural information being processed. So in this sense *neuronal observation is a nonmeasurable reflection (what we have already called a mirroring) of the information that redounds directly from the observer aspect of the measurement process* (see Figure 2.1).

5. Renormalization in a lattice brain. There are two explicit results developed in this section. The first is a notion of observation for a cell assembly along with an expression for its token. The second is that this token increases in value as a hierarchy of assemblies is ascended, a feature that informs a key neuronal competition process developed in Section 7.4.

The neocortex is organized in levels and columns. However since the neural connectivity is irregular, we do not attribute regularity to the neocortical structure. There is literature employing statistical physics for the study of gases under the name Lattice Gases (cui.unige.ch/spc/Bibliography/lattice-gas). There, models of a gas are analyzed imposing a lattice structure on the gas particles. In this section (only), we shall follow an analogous approach. This is done to draw on the development of renormalization methodology in the physics of lattices and to facilitate developments of complete generality (concerning structural geometry and neural connectivity) that follow in Sect. 6.

Consider an assembly of B neurons, the i -th neuron of which has observation token \mathcal{I}_i . (See Section 4.2.) We now specify a notion of observation associated with the assembly, one that arises from ramification of observations of the B individual neurons. We shall adapt from the analysis of the Ising model, this time employing the renormalization group technique. The latter is a coarse graining transformation endowing a block level magnetic structure to blocks of spins induced by the original individual spin interactions. It produces a dynamics whose fixed points specify the transition points of phase change.

We proceed by translating modeling in the special case of spins with nearest neighbor interactions and uniform connection coefficients to our neural context.

As noted, for ease of exposition, we make some temporary specializations of neural geometry and connectivity. The following material is adapted from the discussion of the renormalization group theory applied to Ising model spins in chap. 9 of [14], where additional details such as the specification of notation may also be found. The omitted details are inessential for our argument.

We suppose that N neurons are regularly spaced in a d -dimensional ($d = 1, 2, 3$) lattice with lattice spacing l . We suppose that the assembly is a block of neurons with block spacing that is a multiple of the original lattice spacing. Then each assembly contains l^d neurons. The total number of assemblies is Nl^{-d} . Define a (virtual) output for a neural assembly by the majority rule (a block output)

$$S = \text{sgn} \sum_{i=1}^B S_i. \quad (5.1)$$

The argument proceeds in terms of fixed points of the renormalization group transformation. Then let T^* be the value of the temperature at such a fixed point, and let

$$\tau = \frac{T - T^*}{T^*}. \quad (5.2)$$

(Recall that T is a parameter measuring the noise level.) Let τ_l denote the corresponding normalized temperature associated with the assemblies.

h_l , the exogenous input for the assemblies has the form

$$h_l = hl^{y_h}. \quad (5.3)$$

Let r_i denote the location in space of neuron i . (Recall we are treating the lattice-brain case.)

The renormalization theory specifies the spin correlation as a function $G(|r_i - r_j|, \tau, h)$. (The form of G is inessential for our discussion.) Then referring to (4.2), (4.3) and footnote 2, we have for the observation σ_{ij} that

$$\sigma_{ij} = G(|r_i - r_j|, \tau, h). \quad (5.4)$$

(Recall we identify observation, a process, with its token, a quantity.) It may be shown that

$$G\left(\frac{|r_i - r_j|}{l}, \tau_l, hl^{y_h}\right) = \theta G(|r_i - r_j|, \tau, h), \quad (5.5)$$

where

$$\theta = (l^d / l^{y_h})^2. \quad (5.6)$$

The left side of (5.5) is the block level spin correlation, i.e., it is the assembly level σ .

5.1. Magnification of observation (token) strength in ascending a hierarchy. The preceding development allows us to illustrate a key property of magnification of the value of the token of observation that accompanies progression up the ramification hierarchy via renormalization. We do this by means of an example in the context of the lattice brain. Take the case of dimension $d = 2$ and a triangular lattice (with lattice spacing $l = \sqrt{3}$ and three spins per block). In this case it is known that $y_h = \frac{15}{8}$ (see Chap. 9

of [14]). Then (5.6) gives $\theta = 3^{1/16}$. This shows that the block observation (neuronal assembly observation) is stronger by a factor of 1.15 approximately than the observation at the underlying neuronal level. Iterating the coarsening (ramification) process would successively increase the absolute value of observation by this factor. Indeed every $16/\log_2 3 \approx 10$ iterations doubles the observation magnitude. Since there are $O(10^{11})$ neurons in the human brain, there is capacity for considerable increase of observation magnitude. We shall see in Section 7 that this magnification of observation strength plays a critical role in a neuronal assembly observation competition process.

6. The renormalization of neuronal assemblies. We address the limitations associated with the lattice brain of Section 5 by deriving the renormalization group transformation for a neural assembly, eliminating the restrictions of that model case. That is, we allow arbitrary connections among neurons, both with respect to synaptic (spatial) connectivity and synaptic weight. The regular blocks of Section 5 become subassemblies of neurons within the original neural network, each subassembly being of arbitrary size and spatial configuration. Moreover the connectivity of assemblies, formerly the blocks, is arbitrarily specified by the neuronal connectivity itself (forward, backward, lateral...). We derive the renormalization transformation as well as the key results needed for the application to consciousness in Section 7, namely expressions ((6.24), (6.25)) for the subassembly observation/observation-tokens (the generalization of the block spin correlation function (see (4.2) and footnote 2)). Hereafter subassemblies of the original net are referred to simply as assemblies.

6.1. Derivation of the renormalization transformation. The derivation of the general case ([35]) is an extension of the derivation in [14] in the special Ising model case (with fixed regular lattice geometry). The key is to notice that the synaptic weights w_{ij} (that correspond to the interaction amplitudes in the Ising model case) depend on their indices independently of the method for assigning those indices. To proceed, partition the network into an arbitrary collection of disjoint neural assemblies indexed by the symbol I . Let

$$\{C_I\} = \{S_1^I, S_2^I, \dots\} \quad (6.1)$$

be the collection of possible outputs of the individual neurons in the I -th assembly. Let

$$S_I = \text{sgn}\{S_1^I + S_2^I + \dots\} \quad (6.2)$$

be the virtual output of assembly I .³ Let $H'\{S_I\}$ be the effective energy of the collection of assemblies. Then with $\sum_{\{C_I\}}$ denoting the sum over those C_I that deliver a given one of the values ± 1 of S_I , we write the partition function of the neural net as

$$e^{H'\{S_I\}} = \sum_{\{C_I\}} e^{H\{S_I, C_I\}}. \quad (6.3)$$

For clarity we confine our attention to the case $h^{ex} = 0$. To estimate $H'\{S_I\}$ let

$$H = H_0 + V \quad (6.4)$$

³ S_I is the analog of the block spin of the I -th block in the Ising model. See (4.5)f.

where

$$\begin{aligned} H_0 &= \sum_I \sum_{i,j \in I} w_{ij}^{II} S_i^I S_j^I \\ &\doteq \sum_I S^I W^{II} S^I \end{aligned} \quad (6.5)$$

is the intra-assembly energy (the free part of the Hamiltonian), and where

$$\begin{aligned} V &= \sum_{I \neq J} \sum_{i \in I, j \in J} w_{ij}^{IJ} S_i^I S_j^J \\ &\doteq \sum_{I \neq J} S^I W^{IJ} S^J \end{aligned} \quad (6.6)$$

is the inter-assembly energy (the interactive part of the Hamiltonian). S^I is the vector (S_1^I, S_2^I, \dots) composed of an element of the collection $\{C_I\}$ (that is, a vector whose components are ± 1), and W^{IJ} is the matrix of those synaptic weights connecting neuron j in assembly J to neuron i in assembly I . Now for any function $A(S_I, C_I)$, define the following weighted average:

$$\langle A(S_I) \rangle_0 \doteq \frac{\sum_{\{C_I\}} e^{H_0\{S_I, C_I\}} A(S_I, C_I)}{\sum_{\{C_I\}} e^{H_0\{S_I, C_I\}}}. \quad (6.7)$$

With this, we may rewrite (6.3) as

$$e^{H'\{S_I\}} = \langle e^V \rangle_0 \sum_{\{C_I\}} e^{H_0\{S_I, C_I\}}. \quad (6.8)$$

Next let

$$Z_0(W^{II}) = Z_0(W^{II}, S_I) = \sum_{\{C_I\}} \exp S^I W^{II} S^I. \quad (6.9)$$

Then the free partition function of the entire neural network factors into the product of the free partition functions of each assembly, and we may write

$$\sum_{\{C_I\}} e^{H_0\{S_I, C_I\}} = \prod_I Z_0(W^{II}). \quad (6.10)$$

Combining (6.8) and (6.10) yields

$$e^{H'\{S_I\}} = \langle e^V \rangle_0 \prod_I Z_0(W^{II}). \quad (6.11)$$

The cumulant expansion (see Section 9.6.3 of [14]) gives

$$\langle e^V \rangle_0 = \exp[\langle V \rangle_0 + \frac{1}{2}[\langle V^2 \rangle_0 - \langle V \rangle_0^2] + O(V^3)]. \quad (6.12)$$

Inserting the leading term in this expansion into (6.10), we obtain the approximation

$$H'\{S_I\} = \log \left(\prod_I Z_0(W^{II}) \right) + \langle V \rangle_0. \quad (6.13)$$

Next using (6.6) and the linearity of the averaging operator defined in (6.7), we have

$$\langle V \rangle_0 = \sum_{I \neq J} \langle S^I W^{IJ} S^J \rangle_0. \quad (6.14)$$

Since H_0 does not couple different assemblies the averaging in (6.14) factorizes, giving

$$\langle V \rangle_0 = \sum_{I \neq J} \langle S^I \rangle_0 W^{IJ} \langle S^J \rangle_0. \quad (6.15)$$

Now inserting (6.5) and (6.10) into the averaging formula (6.7), $\langle S^I \rangle_0$ can be written as

$$\langle S^I \rangle_0 = \frac{\sum_{\{C_I\}} e^{S^I W^{II} S^I}}{Z_0(W^{II})}. \quad (6.16)$$

Since $S_I S_I = 1, \forall I$, we may rewrite (6.15) as

$$\langle V \rangle_0 = \sum_{I \neq J} S_I [S_I \langle S^I \rangle_0 W^{IJ} \langle S^J \rangle_0 S_J] S_J. \quad (6.17)$$

Let W_{IJ} denote the assembly level coupling coefficient between assemblies I and J . Then the *renormalization transformation* is obtained by equating W_{IJ} , the coefficient of the interaction $S_I S_J$ energy of the assembly level network, to the bracketed expression in (6.17). Namely

$$W_{IJ} = S_I \langle S^I \rangle_0 W^{IJ} \langle S^J \rangle_0 S_J. \quad (6.18)$$

A more symmetric expression of this relation is

$$S_I W_{IJ} S_J = \langle S^I \rangle_0 W^{IJ} \langle S^J \rangle_0. \quad (6.19)$$

6.2. Observation (tokens) at the assembly level. In this section we shall derive approximate expressions for the tokens of awareness. Employing assembly level analogs of the energy and the neural net partition function (see (6.3)f), we find

$$\langle S_I S_J \rangle = \frac{\partial}{\partial W^{IJ}} \log Z = \frac{\partial}{\partial W^{IJ}} H' \{S_I\}. \quad (6.20)$$

Then using (6.11),

$$\langle S_I S_J \rangle = \frac{\partial}{\partial W^{IJ}} \left[\log \prod_I Z_0(W^{II}) + \log \langle e^V \rangle_0 \right]. \quad (6.21)$$

For $I \neq J$ the first term in the bracket in (6.21) vanishes upon differentiation. Employing the cumulant expansion to its leading term, (6.21) becomes

$$\langle S_I S_J \rangle = \frac{\partial}{\partial W^{IJ}} \log \langle e^V \rangle_0 = \frac{\partial}{\partial W^{IJ}} \langle V \rangle_0, \quad I \neq J. \quad (6.22)$$

Then using (6.15),

$$\langle S_I S_J \rangle = \frac{\partial}{\partial W^{IJ}} \sum_{I \neq J} \langle S^I \rangle_0 W^{IJ} \langle S^J \rangle_0 = \langle S^I \rangle_0 \langle S^J \rangle_0. \quad (6.23)$$

Referring to (4.2) and (4.3), the assembly level observation/observation-tokens are

$$\sigma_{IJ} = \langle S^I \rangle_0 \langle S^J \rangle_0 - \langle S_I \rangle \langle S_J \rangle, \quad I \neq J, \quad (6.24)$$

and since $\langle S_I S_I \rangle = 1$,

$$\sigma_{II} = 1 - \langle S_I \rangle^2. \quad (6.25)$$

The results expressed in these last two equations are closed form expressions for observation tokens and may be expected to play a central role in the development of our understanding of consciousness. Indeed such a theory, developed in the following section, draws on these formalisms.

7. Application: A quantitative theory of consciousness.

7.1. Preliminaries. We develop a quantitative theory of consciousness whose predictions are susceptible to measurement and experimental verification. We employ the model of information processed in a neuronal hierarchy and its renormalization developed in Sections 2–6 to characterize aspects of phenomenal consciousness, focusing on two of its features. These are (i) the so-called explanatory gap that separates the phenomenal experience of first person consciousness and objective prediction/measurement protocols of third person scientific methodology and (ii) the quality of consciousness as an emergent phenomenon.

(i) Explanatory gap

There are many gaps between experience and quantitative scientific methodology. A method for bridging such a gap is to formulate a dualist construct. (See [29], [30] for critical assessments of these approaches.) Examples of such dualist constructs are (a) time and space, (b) unmediated action at a distance as exemplified by Newton’s law of gravity, (c) the Schrödinger wave function, a construct having no existence in reality [10]), which can neither be measured nor observed. While such well-known dualist constructs of science supply us with profound predicting capability, they typically replace one mystery by another. This brings with it a regress: moving the gap into the construct ([47]).

(ii) Emergence of consciousness as a phenomenon of phase change of information

Consciousness is often cited as an emergent phenomenon, somehow emanating from the complex processing of information in the brain. (See [13] for a critical account of this subject.) Our model in Sections 2–6 of the processing of signals (i.e., of information itself) in brain circuitry points to such an emergent characterization by extending the renormalization theory of physics to the information processed in that brain circuitry.

7.2. Measurement, the dualist construct. Our dualist construct is framed in terms of the basic notion of measurement in science, the groundwork of developments in Sections 2–6. It identifies a basic form of awareness, a synaptic level proto-consciousness. This awareness primitive is the reinterpretation of a feature of neuronal processing (for synaptic weight change), namely, Hebb’s law. (See [35] for a motivating discussion. Also see the classical works of [17], [18], [19], [20], [21] that bridge the gap between the facts of neurology and those of psychology.) The reinterpretation flows from viewing neuronal processing as a measurement process as in Section 2.1. There we noted that measurement has two components: (a) a measuring device and (b) an observer of the action of the device. The unconscious processing/transmission of signals in neural circuitry specified in (2.1) embodies both the quantity to be measured (the signal) and the measuring device (the neuron). The Hebbian synaptic dynamics of (2.2) take on the role of the observer.

Together these two dynamical systems specify adjustments Δw_{ij} to the neuron's synaptic weights, adjustments reflecting correlation between (in our terminology, observation of) the input and output of the measuring device. The feature of the Hebbian observer aspect of the neuronal measurement process is taken to be a basic *synaptic awareness* that we call *synaptic consciousness*, and we append it as a dualist construct for our model and its application. We say that the synapse experiences the neural input and output, and the nature of the experience is to cause a change in the value of the synaptic strength.

Appending a dualist construct for the study of consciousness is a well-known approach. Descartes (1637) was one of the first to try this, but see also [4], [35], [39], and [45]. There are many other approaches to the subject ([5], [6], [7], [9], [27], [32], [43], [44]).

7.3. Tokens, experimental verification. The dualist construct enables analytic development, employing tokens, that supports a third person study of consciousness. Tokens of observation developed in Sections 2–6 are reinterpreted as tokens of awareness or consciousness. The formulas that express the tokens in terms of measurable quantities are the basis of experimental verification of the theory. Indeed *measuring that there is adequate strength in a token's value makes for the theory's prediction of a consciousness feature corresponding to that token's awareness counterpart. This prediction could be compared to the experience of the subject.*

Terminology: Although Δw_{ij} is only a *token of the synaptic level of awareness* that itself is the entire (dualist) process represented by the Hebbian dynamics (2.3), we refer to it as the synaptic awareness (consciousness) itself, since confusion should not occur. So Hebbian dynamics (which might be expressible in terms of biochemical processes ([26]) are interpreted as a primitive awareness (a proto-consciousness).

7.4. Neuronal competition. The significance of the magnification feature of observation tokens (developed in Section 5.1) is suggested by a process of neuronal competition. To see this, consider the way in which the preference or choice may be made (for the emergence of the single conscious experience of our awareness) from among the many different signal collections that are being processed simultaneously in the brain (in different hierarchies of neural assemblies). One possibility is to imagine a winner-takes-all competition for coming into consciousness, the winner corresponding to the largest token of awareness. (Competition in neural circuitry is sometimes modeled using lateral inhibition. [34] modified Hebb's theory to include lateral inhibition.) Thus both the increase of the complexity of awareness per step of ascent in the hierarchy and the associated increase in token strength inform the view that *renormalization of information is a route from the dualist construct, the basic synaptic awareness, toward the phenomenal consciousness we experience.*

7.5. Renormalization, ramification. The basic subcellular synaptic consciousness, as a proto-consciousness bears only the most modest resemblance to what is experienced personally as consciousness. We have described the latter to arise out of a sequence of ramification steps of the measurement process. This sequence is constructed by the model, using a *finite hierarchy* of layers of large numbers of inter-communicating neural circuits and their assemblies. (Compare the dynamic core concept of [9].) This agglomerative construction is a recurrent process of renormalization of neuronal information. Experimental evidence for the agglomerative organization of information in assemblies

in the brain and its connection to mood and emotion is found in [38] and in [42]. In the first of these references, the authors show the existence of *ensemble encoding* in the nucleus accumbens, a convergence point in the brain of emotional and cognitive information guiding motivated behavior. In the second, the authors deal with the temporal lobe association cortex organizing information in its pyramidal cells. They identify modification allowing mood states to influence cognitive processing in the brain by facilitating perceptual representations that are *agglomerated organizations of information*.

We identify the coarsening into circuits of neuronal assemblies as the vehicle of ramification of the neural processing/Hebbian dynamics measurement process from its cellular level (i.e., its synapse-neuron level) to the neuronal assembly level. (Recall the notion of mirroring developed in Section 4.4 that motivates this identification.) Iteration of this agglomeration technique is the route from synaptic awareness arising at the cellular level toward the experience familiar to us as consciousness. [28] develops the biological agglomeration case from a primitive sensorium at the cellular level towards circuits specialized in the sensory function.

7.6. Causality or epiphenomenalism? If synaptic weight change is explainable by natural laws in terms of the variables $S_i(n)$ and $S_i(n-1)$ in (2.3), then awareness has epiphenomenal aspects. (Such aspects of consciousness are hardly superficial. They play a role in both daily existence and in evolution [24]). If the weight change occurs because the correlation between these two variables is based on an unmediated process (the dualist construct), then the physical matter that records the synaptic weight is being acted upon by an unmediated phenomenon (a feature no more cryptic than the unmediated attraction at a distance of gravity, say). In this case the *awareness is causal*.

7.7. Influence of noise on awareness, retention of awareness. Recall from (3.3) that $m = m(h)$. From Figure 3.1, we see that in the absence of an exogenous input (i.e., when $h = 0$), we have $m = 0$ for $T \geq T_c$, and so according to (4.8), there is no neuronal awareness in this case. On the other hand, such awareness does return and grow stronger even at above critical temperatures (higher noise levels) with the strengthening of the exogenous input h . Should the input be removed, a generated awareness is preserved by the neuron only if the temperature is below critical (a phase change aspect). So if the noise level is below critical, the neural net can retain the awareness of a retrieved memory after the trigger for that retrieval has been removed.

7.8. Speculations: anomalous phases of information, field of consciousness.

Anomalous states: There are many phases of the information (so-called memory traces) stored in the neural network, and these are characterizable by the renormalization. An identification of these phases with qualia (colors, tastes, sounds, pains...) and feelings (love, hunger, fear, joy...) is suggested.

1. The fixed-point structure of the renormalization group transformation in the more general neural network case is likely to be richer than in the special lattice case of the Ising model exemplified in Section 5. This in turn could lead to an associated phase diagram for the neural network that is more complex (one having more phases...) than the lattice case. We call these new phases *anomalous phases of information*, and we speculate that along with the phase diagram they will help to characterize those features of consciousness such as qualia and feelings that are apparent to our experience.

2. The block structure (a virtual neural network of cell assemblies) is the framework over which the features of ramified awareness levels are supported. According to the renormalization group theory, the virtual net shares the same critical point/phase diagram structure as the original neurons. Thus the anomalous phases and the associated phase transition diagram are available to the ramified awareness levels (ultimately, to the phenomenal consciousness of our experience). Likewise the tokens of our actual consciousness should have the form shown in (6.24), (6.25).

Field of consciousness: A collection of magnetic spins may change phase, taking form from among ferromagnetic, anti-ferromagnetic, or disordered states. In the Hopfield model, the set of available phases generalizes to the rich collection of possible memory traces of the neural net. Among these traces are the fundamental memories (see (3.3)ff) and the so-called *spurious states* (see [35]). We defer the production and interpretation of associated phase diagrams to future work.

Except for the critical inclusion of the Hebbian dynamics, the neural modeling discussed in Sections 2–4 is a variant of the mean field analysis of the Ising model of magnetization. (See [22] Chap. 2.) The quantity \bar{S} in (4.5) is the magnetization M of the system in the corresponding Ising model (see eqn.(3.19) or eqn.(4.32) in [1]). This analogy describes the awareness (defined here) associated with the neurons in a network as corresponding to the magnetization surrounding the dipoles composing a magnet. Thus our model prompts the speculation that *consciousness is a field of awareness associated with a neural network*. In what sense this field might surround the network (as the magnetic field surrounds the magnet), we cannot say. Note the relationship of this comment to the fundamental Hopfield-like quantum neural holography approach to consciousness of [40] and of [41].

REFERENCES

- [1] Amit, D. (1989), *Modeling Brain Function, The World of Attractor Neural Networks*, Cambridge Univ. Press. MR1025121 (90k:92003)
- [2] Amit, D., Gutfreund, H., Sompolinsky, H. (1987), *Statistical Mechanics of Neural Networks near Saturation*, *Annals of Physics* **173**, 30-67.
- [3] Beckermann, M. (1997), *Adaptive Cooperative Systems*, Wiley, New York.
- [4] Chalmers, D. (1996), *The Conscious Mind: In Search of a Fundamental Theory*, Oxford Univ. Press.
- [5] Churchland, P. M. (1984), *Matter and Consciousness*, MIT Press, Cambridge.
- [6] Damasio, A. (1994), *Descartes Error*, Grosset/Putnam, New York.
- [7] Dennett, D. (1996), *Kinds of Minds, Toward an Understanding of Consciousness*, Basic Books, New York.
- [8] Descartes, R. (1637), *Discours sur la Méthode*.
- [9] Edelman, G., Tononi, G. (2001), *A Universe of Consciousness: How Matter Becomes Imagination*, Basic Books, New York.
- [10] Einstein, A., Podolsky, B., Rosen, N. (1935), *Can Quantum-Mechanical Description of Physical Reality be Considered Complete?*, *Phys. Rev.* **47**, 777-780.
- [11] Franssen, E., Lansner, A. (1995), *Low spiking rates in a population of mutually exciting pyramidal cells*, *Network* **6**, 271-288.
- [12] ——— (1998), *A model of cortical associative memory based on a horizontal network of connected columns*, *Network* **9**, 235-264.
- [13] Freeman, A., Ed. (2001), *The Emergence of Consciousness*, *J. Consc. Studies* **8**, 9-10.
- [14] Goldenfeld, N. (1992), *Lectures on Phase Transitions and the Renormalization Group*, Addison-Wesley, Reading, MA.

- [15] Haken, H. (2004), *Synergetic Computers and Cognition*, Springer, New York. MR2081217 (2005e:68207)
- [16] Haykin, S. (1999), *Neural Networks, A Comprehensive Foundation*, Prentice Hall, Upper Saddle River, NJ.
- [17] Hebb, D. (1939), *Intelligence in man after large removals of cerebral tissue: Report of four left frontal lobe cases*, J. Gen. Psychol., 73-87.
- [18] ——— (1946), *On the nature of fear*, Physiol. Rev. **53**, 259-276.
- [19] ——— (1949), *Organization of Behavior: A Neurophysiological Theory*, Wiley, NY.
- [20] ——— (1980a), *The structure of thought*, in *The Nature of Thought*, P.Juszyk and R.Klein, Eds., Lawrence Erlbaum Associates, Hillsdale, NJ.
- [21] ——— (1980b), *Essay on Mind*, Lawrence Erlbaum Associates, Hillsdale, NJ.
- [22] Hertz, J., Krogh, A., Palmer, R. (1991), *Introduction to the Theory of Neural Computation*, Addison-Wesley, Redwood City, CA. MR1096298 (92f:82044)
- [23] Hopfield, J. (1982), *Neural networks and physical systems with emergent collective computation abilities*, PNAS, USA **79**, 2254-2258. MR652033 (83g:92024)
- [24] Johnston, V. (1999), *Why We Feel, The Science of Human Emotion*, Perseus, Cambridge, MA.
- [25] Kadanoff, L. (1966), *Scaling laws for Ising models near T_c* , *Physics*, vol. 2, pp. 263-272.
- [26] Kandel, E.R., Schwartz, J.H., Jessel, T.M. (2000), *Principles of Neural Science*, 4th ed., McGraw-Hill/Appleton & Lange, New York.
- [27] Libet, B. (2003), *Neurophysiology of Consciousness: Selected Papers and New Essays by Benjamin Libet*, Birkhäuser, Boston.
- [28] Llinas, R. (2001), *i of the Vortex*, MIT Press, Cambridge.
- [29] Margenau, H. (1950), *The Nature of Physical Reality: A Philosophy of Modern Physics*, McGraw-Hill, New York. MR0035258 (11:709h)
- [30] ——— (1978), *Physics and Philosophy: Selected Essays*, Dordrecht, Boston. MR0529767 (80c:01020)
- [31] McCullough, W., Pitts, W. (1943), *A logical calculus of the ideas immanent in nervous activity*, Bull. Math. Biophysics **5**, 115-133. MR0010388 (6:12a)
- [32] McDermott, D. (2001), *Mind and Mechanism*, MIT Press, Cambridge.
- [33] Mezard, M., Parisi, G., Virasoro, M. (1987), *Spin Glass Theory and Beyond*, Singapore, World Scientific. MR1026102 (91k:82066)
- [34] Milner, P. (1957), *The cell assembly: Mark II*, Psychol. Rev. **64**, 242-252.
- [35] Miranker, W.L. (2000), *Consciousness is an information state*, J. Neural Parallel and Scientific Computation **8**, 83-104.
- [36] ——— (2001), *The Renormalization of Information*, Yale Univ. DCS/TR-1215.
- [37] ——— (2005), *Consciousness, A Darwinian Process*, Yale Univ. DCS/TR-1344.
- [38] O'Donnell, P., Greene, J., Pabello, N., Lewis, B., Grace, A. (1999), *Modulation of cell firing in the nucleus accumbens*, Annals of the NY Academy of Sciences **877**, 157-175.
- [39] Penrose, R. (1994), *Shadows of the Mind*, Oxford University Press, Oxford. MR1865778 (2002g:00003)
- [40] Perus, M. (2001), *Multi-level synergetic computation in brain*, Nonlinear Phenomena in Complex Systems **4**, 157-193; cf. also www.iasc-bg.org.yu. MR1886860 (2003e:92003)
- [41] Rakovi, D., Dugi, M. (2005), *Quantum-holographic classical Hopfield-like associative nets: Implications for modeling two cognitive modes of consciousness*, Opticheski J. **75**, 13-18 (Special issue on Topical Meeting on Optoinformatics 'Optics Meets —' Saint Petersburg (2004); also www.iasc-bg.org.yu.
- [42] Rolls, E., Treves, A. (1999), *Neural Networks and Brain Function*, Oxford Univ. Press, Oxford.
- [43] Searle, J. (1994), *The Rediscovery of the Mind*, MIT Press, Cambridge.
- [44] Sheets-Johnstone, M. (1998), *Consciousness; a natural history*, J. Consc. Studies **3**, 260-294.
- [45] Stapp, H. (1998), *The Hard problem: a quantum approach*, J. Consc. Studies **3**, 194-210.
- [46] Thompson, M., Varela, F. (2001), *Radical embodiments: neural dynamics and consciousness experiences*, Trends in Cognitive Sciences **5**, 418-425.
- [47] Velmans, M. (2000), *Understanding Consciousness*, Routledge, London.
- [48] Wilson, K. (1971), *Renormalization group and critical phenomena I & II*, Phys. Rev. **B4**, 3174-3183 and 3184-3205.